Math 703: Problem Set 2

Huzaifa Mustafa Unjhawala

Due: October 5, 11:59 PM

1. It is usually helpful for a numerical scheme to preserve important symmetries and conservation principles associated with the mathematical statement of an application problem. This problem demonstrates that not all numerical schemes are equivalent for preservation of symmetry. Starting from the identity

$$-\frac{d}{dx}\left[C(x)\frac{du(x)}{dx}\right] = -C(x)\frac{d^2u(x)}{dx^2} - \frac{dC(x)}{dx}\frac{du(x)}{dx}, \tag{1}$$

assume that $C(x)$ is known and can be evaluated at all values of x, while $u(x)$ is only known at the grid points x_i.

(a) Find a centered difference approximation to the LHS of (1) that is symmetric.

(b) Show that the centered difference approximation to the RHS is not exactly symmetric, and show that the approximation is only symmetric up to an error of $O(\delta x^3)$.

Solution:

(a) Taking first the centered difference approximation of $\frac{du}{dx}$, we get

$$\frac{du}{dx} \approx \frac{u_{i+1} - u_{i-1}}{2\delta x}$$

Here, note that $u_{i+1} = u(x_{i+1})$ and $u_{i-1} = u(x_{i-1})$. Then, multiplying by $C(x)$ on both sides, we get

$$\left[C(x)\frac{du}{dx}\right] \approx \left[C(x)\frac{u_{i+1} - u_{i-1}}{2\delta x}\right]$$

Now, taking the centered difference approximation of $\left[C(x)\frac{u_{i+1}-u_{i-1}}{2\delta x}\right]$, we get

$$\frac{d}{dx}\left[C(x)\frac{u_{i+1} - u_{i-1}}{2\delta x}\right] \approx \frac{C(x_{i+1})\frac{u_{i+2}-u_i}{2\delta x} - C(x_{i-1})\frac{u_i-u_{i-2}}{2\delta x}}{\delta x}$$

Expanding the numerator of the RHS, we get

$$\frac{C(x_{i+1})(u_{i+2}-u_i)}{2\delta x} - \frac{C(x_{i-1})(u_i-u_{i-2})}{2\delta x}$$

Now, we can take the negative of the RHS to get the LHS of (1).

$$-\frac{d}{dx}\left[C(x)\frac{u_{i+1}-u_{i-1}}{2\delta x}\right] = -\frac{C(x_{i+1})(u_{i+2}-u_i)}{2\delta x} + \frac{C(x_{i-1})(u_i-u_{i-2})}{2\delta x}$$

Putting back the denominator, we get

$$-\frac{d}{dx}\left[C(x)\frac{u_{i+1}-u_{i-1}}{2\delta x}\right] \approx -\frac{C(x_{i+1})(u_{i+2}-u_i) - C(x_{i-1})(u_i-u_{i-2})}{4\delta x^2}$$

As we can see, the LHS is symmetric about x_i and since we have a second derivative, we need terms such as x_{i-2} and x_{i+2} to be included in the approximation.

(b) Again, taking the centered difference of the different terms at x_i, we get

$$-C(x)\frac{d^2u(x)}{dx^2} \approx -C(x)\frac{u_{i+1}-2u_i+u_{i-1}}{\delta x^2}$$

And the other term is

$$\frac{dC}{dx}\frac{du}{dx} \approx \frac{C(x_{i+1})-C(x_{i-1})}{2\delta x}\cdot\frac{u_{i+1}-u_{i-1}}{2\delta x}$$

Expanding this, we get

$$\frac{dC}{dx}\frac{du}{dx} \approx \frac{(C(x_{i+1})-C(x_{i-1}))(u_{i+1}-u_{i-1})}{4\delta x^2}$$

Further expanding the numerator, we get

$$\frac{dC}{dx}\frac{du}{dx} \approx \frac{C(x_{i+1})u_{i+1} - C(x_{i+1})u_{i-1} - C(x_{i-1})u_{i+1} + C(x_{i-1})u_{i-1}}{4\delta x^2}$$

As we can see, there are unsymmetric terms in this second term such as $C(x_{i+1})u_{i-1}$ and $C(x_{i-1})u_{i+1}$. Therefore, the RHS is not symmetric about x_i. These asymmetries arise due to the centered difference approximation of $\frac{dC}{dx}$.

2. *Note: I suggest grounding as the last step.*

(a) For the circuit drawn in lecture, and with resistors on every edge, show that the topology matrix A has dependent columns. Then ground one of the nodes, and write down the reduced system of equations. Show that the reduced $\tilde{A}$ has independent columns, and hence conclude that $\tilde{A}^T\tilde{C}\tilde{A}$ is symmetric and positive definite.

(b) Put a current source on one edge. How does the system of equations change? Then ground, and check to see if the reduced system is associated with a symmetric, positive definite matrix.

Solution:

(a) From class, from the circuit drawn, we have the incidence matrix as

$$\begin{pmatrix} -1 & 0 & 0 & 0 & +1 & 0 & 0 & 0 \\ +1 & -1 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & +1 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & +1 & -1 & 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 & -1 & +1 & 0 & +1 \end{pmatrix}$$

Taking the transpose of this matrix, we get

$$\begin{pmatrix} -1 & +1 & 0 & 0 & 0 \\ 0 & -1 & +1 & 0 & 0 \\ 0 & 0 & -1 & +1 & 0 \\ 0 & 0 & 0 & -1 & +1 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Now, to see if the columns are dependent, we just add all the columns. We get the zero vector. Therefore, the columns are dependent. Now, lets assume that we ground the 5th node (indexed 1). Then we remove the 5th column of the matrix and get the new modified matrix as

$$\begin{pmatrix} -1 & +1 & 0 & 0 \\ 0 & -1 & +1 & 0 \\ 0 & 0 & -1 & +1 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

Now we can see that we cannot write any column vector as a linear combination of the other column vectors. Therefore, the columns are independent. Matrix $\mathbf{C}$ is a idenitity matrix since the resistors are all 1. Therefore, $\mathbf{C} = \mathbf{I}$. Therefore, $\tilde{\mathbf{A}}^T\mathbf{C}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}^T\mathbf{I}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}^T\tilde{\mathbf{A}}$ and this evaluates to

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & -1 & -1 & 3 \end{pmatrix}$$

This is clearly symmetric. I then used Python to compute the eigen values of this matrix and found them to be $5, 2, 1, 4$. These are all positive and real, thus the matrix is positive definite.

(b) Lets put a current source on edge 8 (indexed 1). Let this current source be of value 1. Now, we would need to take the 8th column of the $\mathbf{A}^T$ and multiply it by 1 and take it to the RHS. Thus, we have $\tilde{\mathbf{A}}^T$ becomes

$$\begin{pmatrix} -1 & 0 & 0 & 0 & +1 & 0 & 0 \\ +1 & -1 & 0 & 0 & 0 & -1 & -1 \\ 0 & +1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & -1 & 0 & 0 & +1 \\ 0 & 0 & 0 & +1 & -1 & +1 & 0 \end{pmatrix}$$

Then, we also remove the corresponding y from the y vector and the equations become

$$\tilde{\mathbf{A}}^T \tilde{\mathbf{y}} = \tilde{\mathbf{F}}$$

Where $\tilde{\mathbf{A}}^T$ is given as above, $\tilde{\mathbf{y}}$ is just the y's without y_8 and $\tilde{\mathbf{F}}$ is

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

Now, if we ground again the 5th node, we can just remove the 5th row of the matrix and we get

$$\begin{pmatrix} -1 & 0 & 0 & 0 & +1 & 0 & 0 \\ +1 & -1 & 0 & 0 & 0 & -1 & -1 \\ 0 & +1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & -1 & 0 & 0 & +1 \end{pmatrix}$$

Again, if we do $\tilde{\mathbf{A}}^T \mathbf{C} \tilde{\mathbf{A}}$, we will get the same matrix as before and we will see that the matrix is still symmetric and positive definite.

3. Consider minimization of $Q(y) = 0.5y^T C^{-1} y - b^T y$ subject to the constraint $A^T y = f$, and maximization of $-P(x) = -0.5(Ax - b)^T C(Ax - b) - x^T f$. Assuming C is symmetric and positive definite, prove duality starting from the sum $Q + P$.

Solution:

First we compute the sum $Q(\mathbf{y}) + P(\mathbf{x})$

$$Q(\mathbf{y}) + P(\mathbf{x}) = \frac{1}{2}\mathbf{y}^\top \mathbf{C}^{-1}\mathbf{y} - \mathbf{b}^\top \mathbf{y} - \frac{1}{2}(\mathbf{A}\mathbf{x} - \mathbf{b})^\top \mathbf{C}(\mathbf{A}\mathbf{x} - \mathbf{b}) - \mathbf{x}^\top \mathbf{f}$$

To solve the minimization problem we take the lagrangian like in class:

$$L(\mathbf{y}, \mathbf{x}) = \frac{1}{2}\mathbf{y}^\top \mathbf{C}^{-1}\mathbf{y} - \mathbf{b}^\top \mathbf{y} + \mathbf{x}^\top (\mathbf{A}^\top \mathbf{y} - \mathbf{f})$$

The optimality condition requires that the gradient of L with respect to $\mathbf{y}$ is zero:

$$\nabla_{\mathbf{y}} L = \mathbf{C}^{-1}\mathbf{y} - \mathbf{b} + \mathbf{A}\mathbf{x} = 0$$

Solving for $\mathbf{y}$, we get:

$$\mathbf{y} = \mathbf{C}(\mathbf{b} - \mathbf{A}\mathbf{x})$$

From this, it follows that:

$$\mathbf{A}\mathbf{x} - \mathbf{b} = -\mathbf{C}^{-1}\mathbf{y}$$

Next, we simplify the quadratic terms. We have:

$$(\mathbf{A}\mathbf{x} - \mathbf{b})^\top \mathbf{C}(\mathbf{A}\mathbf{x} - \mathbf{b}) = \left(-\mathbf{C}^{-1}\mathbf{y}\right)^\top \mathbf{C}\left(-\mathbf{C}^{-1}\mathbf{y}\right) = \mathbf{y}^\top \mathbf{C}^{-1}\mathbf{y}$$

Thus,

$$\frac{1}{2}\mathbf{y}^\top \mathbf{C}^{-1}\mathbf{y} - \frac{1}{2}(\mathbf{A}\mathbf{x} - \mathbf{b})^\top \mathbf{C}(\mathbf{A}\mathbf{x} - \mathbf{b}) = 0$$

From here, we can simplify the sum $Q(\mathbf{y}) + P(\mathbf{x})$ as follows:

$$Q(\mathbf{y}) + P(\mathbf{x}) = -\mathbf{b}^\top \mathbf{y} - \mathbf{x}^\top \mathbf{f}$$

Since $\mathbf{A}^\top \mathbf{y} = \mathbf{f}$, we can express $\mathbf{x}^\top \mathbf{f}$ as:

$$\mathbf{x}^\top \mathbf{f} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{y} = (\mathbf{A}\mathbf{x})^\top \mathbf{y}$$

Substituting this expression into the equation, we have:

$$Q(\mathbf{y}) + P(\mathbf{x}) = -\mathbf{b}^\top \mathbf{y} - \mathbf{y}^\top \mathbf{A}\mathbf{x} = -\mathbf{y}^\top (\mathbf{A}\mathbf{x} + \mathbf{b})$$

We know that $\mathbf{y} = \mathbf{C}(\mathbf{b} - \mathbf{A}\mathbf{x})$. Therefore,

$$\mathbf{y}^\top (\mathbf{A}\mathbf{x} + \mathbf{b}) = [\mathbf{C}(\mathbf{b} - \mathbf{A}\mathbf{x})]^\top (\mathbf{A}\mathbf{x} + \mathbf{b})$$

Since $\mathbf{b} - \mathbf{A}\mathbf{x}$ and $\mathbf{A}\mathbf{x} + \mathbf{b}$ are vectors, their dot product simplifies to zero due to symmetry, implying that:

$$Q(\mathbf{y}) + P(\mathbf{x}) = 0$$

At the optimal $\mathbf{y}$ and $\mathbf{x}$, the sum $Q(\mathbf{y}) + P(\mathbf{x})$ equals zero, demonstrating the duality between the minimization and maximization problems.

4. For K symmetric and positive definite, minimize $Q(y) = y^T K y$ subject to $y^T y = 1$. Show that the minimum value of $Q(y)$ is the smallest eigenvalue of K. **Solution:**

Since K is symmetric and positive definite, it can be diagonalized by an orthogonal matrix. This means there exists an orthonormal set of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and corresponding positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ such that:

$$\mathbf{K}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad \text{for } i = 1, 2, \ldots, n.$$

Thus, any vector $\mathbf{y}$ satisfying $\mathbf{y}^T\mathbf{y} = 1$ can be expressed as a linear combination of the eigenvectors:

$$\mathbf{y} = \sum_{i=1}^{n} c_i \mathbf{v}_i,$$

where the coefficients c_i satisfy:

$$\mathbf{y}^T\mathbf{y} = \left(\sum_{i=1}^{n} c_i \mathbf{v}_i\right)^T \left(\sum_{j=1}^{n} c_j \mathbf{v}_j\right) = \sum_{i=1}^{n} c_i^2 = 1.$$

Substituting the expression for $\mathbf{y}$ into $Q(\mathbf{y}) = \mathbf{y}^T K \mathbf{y}$, we get:

$$Q(\mathbf{y}) = \left(\sum_{i=1}^{n} c_i \mathbf{v}_i\right)^T K \left(\sum_{j=1}^{n} c_j \mathbf{v}_j\right) \tag{1}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \mathbf{v}_i^T K \mathbf{v}_j \tag{2}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \mathbf{v}_i^T \lambda_j \mathbf{v}_j \tag{3}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \lambda_j \mathbf{v}_i^T \mathbf{v}_j, \tag{4}$$

$$= \sum_{i=1}^{n} c_i^2 \lambda_i. \quad \text{Due to orthonormality of } \mathbf{v}_i. \tag{5}$$

Thus, the problem reduces to minimizing $\sum_{i=1}^{n} c_i^2 \lambda_i$ subject to $\sum_{i=1}^{n} c_i^2 = 1$. Since all λ_i are positive, we should assign the largest weight to the smallest eigenvalue. This in our case is λ_n. Thus taking $c_i = 0$ for all $i \neq n$ and $c_n = 1$, we get the minimum value of $Q(y)$ as λ_n.

5.

(a) Consider n masses and n springs under the action of gravity, with a fixed-end boundary condition at the top wall and a free-end boundary condition at the bottom (at mass n). Find the incidence matrix A and its transpose A^T. (If helpful, use the 3×3 case as an example.)

Now consider an elastic rod of the same length as the above mass-spring system in the gravity-off (rest) configuration. For the elastic rod, we wish to turn on gravity and find the equilibrium ODE for displacement $u(x)$ in the stretched configuration, where we take $n \to \infty$ keeping the rest length fixed.

The discrete equilibrium becomes

$$-\frac{d}{dx}\left(C^*(x)\frac{du(x)}{dx}\right) = f^*(x), \quad u(0) = 0, \quad \left.\frac{du}{dx}\right|_{x=1} = 0, \tag{5}$$

where $f^*(x)$ is force per unit length. Parts (b)-(d) ask you to consider the relations underlying (5), by analogy with the discrete relations leading to $A^TCAx = f$ as discussed in class.

(b) What is the continuous analog of the discrete relation $e = Ax$ (see class notes). Note that this condition includes the fixed-end boundary condition at $x = 0$.

(c) What is the continuous version of the constitutive law $y = Ce$? Explain why the discrete matrix C with spring constants k as diagonal elements must now be replaced by a continuous function $C^*(x)$, where $C^*(x)$ has dimensions of force.

(d) What is the continuous version of the equilibrium condition $A^Ty = f$? Note that f is replaced by the force per unit length $f^*(x)$. Explain the meaning of the boundary condition at $x = 1$.

Solution:

(a) Considering the 3 mass spring system and $\mathbf{A}$ as in the class notes, all we need to do is add a column to the $\mathbf{A}$ we derived in class because now even the last node which was fixed in class is free to move. We will define this end that is free to move as node 3 and its position will be given by x_3. Thus our incidence matrix will be

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

The transpose of this matrix will be

$$\mathbf{A}^T = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

(b) $e = \frac{du(x)}{dx}$ since $\mathbf{A}$ is the difference matrix.

(c) $w(x) = C^*(x)\frac{du(x)}{dx}$. In terms of dimensions, the $\mathbf{e}$ used to be dimensions of length, but now is dimensionless (since $\frac{du(x)}{dx}$ is dimensionless). Thus to match the dimensions of w which has units of force, we multiply by $C^*(x)$ which has dimensions of force. It goes from a discrete function defined between grid points to a continuous function defined for all x as there are an infinite number of grid points in the continuous case.

(d) $\mathbf{A}^T\mathbf{y} = \mathbf{F}$ in continuous case can we written as $-\frac{dw(x)}{dx} = f^*(x)$. However, we can use the relation $\mathbf{y} = \mathbf{CAx}$ and get the continuous version as $-\frac{d}{dx}\left(C^*(x)\frac{du(x)}{dx}\right) = f^*(x)$. The meaning of the boundary condition at $x = 1$ which is $\left.\frac{du}{dx}\right|_{x=1} = 0$ implies that, at $x = 1$ the rod is free of any external forces or constraints. Actually it also implies that the internal force, given in the discrete case by $\mathbf{y}$ and the continuous case by $w(x)$ is zero at $x = 1$. If we consider elasticity, then this means that there is no strain at this point on the rod — hence the name "free-end", as in the truest sense, it is free from strain and thus from stress.